JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 107, NO. D15, 10.1029/2001JD001247, 2002

Comparison of mean climate trends in the Northern Hemisphere between National Centers for Environmental Prediction and two atmosphere-ocean model forced runs

Valerio Lucarini[1]

Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Gary L. Russell

NASA Goddard Institute for Space Studies, New York, New York, USA

Received 23 August 2001; revised 8 December 2001; accepted 26 December 2001; published 13 August 2002.

[1] Results are presented for two greenhouse gas experiments of the Goddard Institute for Space Studies atmosphere-ocean model (AOM). The computed trends of surface pressure; surface temperature; 850, 500, and 200 mbar geopotential heights; and related temperatures of the model for the time frame 1960–2000 are compared with those obtained from the National Centers for Enviromental Prediction (NCEP) observations. The domain of interest is the Northern Hemisphere because of the higher reliability of both the model results and the observations. A spatial correlation analysis and a mean value comparison are performed, showing good agreement in terms of statistical significance for most of the variables considered in the winter and annual means. However, the 850 mbar temperature trends do not show significant positive correlation, and the surface pressure and 850 mbar geopotential height mean trends confidence intervals do not overlap. A brief general discussion about the statistics of trend detection is presented. The accuracy that this AOM has in describing the regional and NH mean climate trends inferred from NCEP through the atmosphere suggests that it may be reliable in forecasting future climate changes. *INDEX TERMS:* 1620 Global Change: Climate dynamics (3309); 1694 Global Change: Instruments and techniques; 3309 Meteorology and Atmospheric Dynamics: Climatology (1620); 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; *KEYWORDS:* historical climate change, Northern Hemisphere climatology, climate trends detection, statistical methods in climate data analysis, comparison with NCEP data, atmospheric data

1. Introduction

[2] A complete quantitative comparison of the results of a coupled model with data coming from the observations is needed to test the validity of the model analyzed. In order to assess the credibility of a model to describe climate change, it is necessary to perform a statistical study that analyzes the compatibility of spatial averages and spatial patterns between model and observed trends of climatologically relevant quantities. In the present report we perform a thorough comparison of the results of two greenhouse gas (GHG) experiments (minus their control simulations) of the Goddard Institute for Space Studies (GISS) atmosphere-ocean model (AOM) developed by *Russell et al.* [1995] with climatological data from National Centers for Environmental Prediction (NCEP) reanalysis (available at http://www.cdc.noaa.gov/cdc/data.ncep.reanalysis.derived.html) in the time frame 1960–2000. We have limited our analysis to the Northern Hemisphere (NH) because in the NH the model is more stable, and the NCEP data, especially the older ones, are generally more reliable than in the Southern Hemisphere (SH). The variables analyzed here are surface pressure; surface temperature; and geopotential height and temperature at 850, 500, and 200 mbar.

[3] The two GHG experiments differ by their initial conditions. The first experiment GHG1 and its control simulation were started after a 40 year spin-up from *Levitus et al.* [1994] ocean conditions with 1950 atmospheric composition. GHG2 and its control were spun up for 100 years from Levitus conditions. The first 60 years (1950–2010) of GHG1's control shows a climate drift of 0.5 K in surface air temperature, whereas the first 90 years of GHG2 shows a drift of <0.1 K. For the first 50 years, there are major disagreements in the SH between the two model runs; for example, Antarctica warms in GHG1 but cools in GHG2. The surface temperature data of these model simulations have already been used by *Russell et al.* [2000] to compare regional changes with observational data compiled by *Hansen et al.* [1996, 1999] for the time period 1960–

[1]Also at Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

0148-0227/02/2001JD001247$09.00

Table 1. NH Spatial Correlation Coefficients of Seasonal and Annual Trends for 1960–2000 for Temperature[a]

Temperature	Experiment	DJF	MAM	JJA	SON	ANN
TAS	NCEP versus GHGs	(0.511)	(0.550)	−0.152	0.406	(0.519)
	NCEP versus GHG1	(0.414)	0.247	−0.173	0.366	(0.383)
	NCEP versus GHG2	(0.435)	(0.573)	−0.046	0.356	(0.451)
	GHG1 versus GHG2	0.372	0.222	−0.008	(0.578)	(0.307)
T850	NCEP versus GHGs	0.364	0.163	−0.225	−0.258	0.042
	NCEP versus GHG1	0.303	0.015	−0.249	−0.301	0.079
	NCEP versus GHG2	0.216	0.208	−0.080	−0.152	−0.005
	GHG1 versus GHG2	−0.006	0.065	0.065	0.446	0.036
T500	NCEP versus GHGs	(0.466)	0.418	0.036	0.189	(0.368)
	NCEP versus GHG1	0.424	0.219	−0.058	0.130	0.263
	NCEP versus GHG2	0.233	0.369	0.115	0.173	0.194
	GHG1 versus GHG2	0.027	−0.027	0.094	0.284	−0.224
T200	NCEP versus GHGs	(0.511)	0.391	0.185	0.335	(0.498)
	NCEP versus GHG1	0.389	0.291	0.127	0.335	(0.429)
	NCEP versus GHG2	(0.488)	0.359	0.215	0.294	(0.483)
	GHG1 versus GHG2	0.452	0.355	(0.667)	(0.692)	(0.682)

[a] Temperature abbreviations are surface air, 850, 500, and 200 mbar temperatures, respectively. Values in parentheses exceed the 0.95 confidence level. DJF is December, January, and February; MAM is March, April, and May; JJA is June, July, and August; SON is September, October, and November; and ANN is annual (as used in text).

1998. We refer to *Russell et al.* [2000] for further description of these model simulations.

[4] We have decided to perform the analysis of the surface temperature trend again because in this study it could be inserted in a broader context. Since the purpose of this work is essentially to perform an analysis of the model's pressure and temperature fields forced changes through the atmosphere, for reasons of self-consistency it is critical to compare them with complete data sets coming from the same source. The NCEP reanalysis [*Kalnay et al.*, 1996] provides data for various variables which are synchronically coherent since they are outputs of a model. Another important feature of these data is that they go relatively far back in time.

[5] There is not a lot of confidence in the diachronic coherence of the NCEP data, and therefore in the deduced trends, essentially because of the changes in the data assimilations which have occurred in the past. There are evidences of shifts with time of the NCEP temperatures with respect to those derived from the analysis of the channel 2 of the microwave sounding unit (MSU) data, even if the spatial pattern of the anomalies usually compares well [*Trenberth et al.*, 2001; *Basist et al.*, 1997; *Chelliah and Ropelewski*, 2000]. A recent study [*Chelliah and Ropelewski*, 2000] shows very good agreement between the spatial patterns of the surface temperature anomalies derived from NCEP and from surface observations for 1958–1998, which roughly corresponds to the time frame considered in this study. Some preliminary studies (V. Lucarini, unpublished results, 2001) suggest that the NCEP surface temperature and 850–300 mbar layer virtual temperature 1960–2000 mean trends for the global, NH, SH, and tropical regions have 0.95 confidence intervals that strongly overlap with those derived from *Angell* [1999] radiosonde data. However, radiosonde [*Gaffen et al.*, 2000] and MSU [*Hurrell and Trenberth*, 1998; *Hurrell et al.*, 2000] data themselves do suffer from biases; they should not be considered as absolutely reliable standards.

2. Procedures

[6] The resolution of the freely available NCEP reanalysis monthly averaged data is 2.5° × 2.5°. The first procedure is to interpolate those data to the model resolution, which is 4° latitude × 5° longitude. We then subtract the corresponding control data set from the data sets of each of the two GHG experiments in order to reduce the effect of climate drift. This processing has been performed in two different ways: first, by subtracting from each greenhouse gas experiment a 21 year moving average of its control run; and second, by subtracting the control runs year by year. The two procedures give very similar but not indistinguishable results; results relative to the the 21 year averaging technique as given by *Russell et al.* [2000] are presented here because they are conceptually closer to the idea of climate drift subtraction. We thus obtain the two model data sets GHG1 and GHG2 for each variable. In order to reduce the influence of model noise, a new data set for all variables is created by averaging GHG1 and GHG2 and is named GHGs.

[7] At every grid cell and for every variable previously described, the trend is computed as the slope of the least squares fit line of the seasonal or annual values for 1960–2000. Then, for each of these variables the area-weighted spatial correlation coefficients between the NH trends of NCEP versus GHG1, NCEP versus GHG2, NCEP versus GHGs, and GHG1 versus GHG2 are computed (see *Johns et al.* [2001] for another spatial correlation analysis of modeled and observed surface temperature trends). The last two correlations are the most relevant ones because GHGs contains the most statistically significant information that we can deduce from the model runs, while comparing GHG1 versus GHG2 gives us an estimate of the model's self-consistency and stability. In all grid cells where the 850 mbar geopotential height has crossed the surface at least once in the 41 year record the 850 mbar data have been discarded from all analyses. These grid cells are blank on the GHGs 850

Table 2. NH Spatial Correlation Coefficients of Seasonal and Annual Trends for 1960–2000 for Surface Pressure[a]

Experiment	DJF	MAM	JJA	SON	ANN
NCEP versus GHGs	(0.663)	(0.458)	0.056	0.172	(0.490)
NCEP versus GHG1	(0.628)	0.315	0.039	0.041	(0.377)
NCEP versus GHG2	0.388	0.210	0.051	0.190	0.281
GHG1 versus GHG2	(0.479)	−0.358	0.249	0.021	0.198

[a] Values in parentheses exceed the 0.95 confidence level.

Table 3. NH Spatial Correlation Coefficients of Seasonal and Annual Trends for 1960–2000 for Geopotential Height[a]

Geopotential Height	Experiment	DJF	MAM	JJA	SON	ANN
Z850	NCEP versus GHGs	(0.600)	0.382	−0.095	0.034	0.264
	NCEP versus GHG1	(0.575)	0.270	−0.068	−0.100	0.227
	NCEP versus GHG2	0.263	0.181	−0.072	0.135	0.063
	GHG1 versus GHG2	0.119	−0.294	0.071	−0.063	−0.320
Z500	NCEP versus GHGs	(0.617)	0.386	0.017	−0.164	0.211
	NCEP versus GHG1	(0.622)	0.215	−0.048	−0.233	0.277
	NCEP versus GHG2	0.218	0.279	0.081	−0.037	−0.046
	GHG1 versus GHG2	−0.004	−0.179	−0.084	0.191	−0.352
Z200	NCEP versus GHGs	(0.748)	(0.469)	0.021	−0.072	(0.564)
	NCEP versus GHG1	(0.757)	0.344	−0.018	−0.091	(0.544)
	NCEP versus GHG2	0.296	0.210	0.060	−0.014	−0.006
	GHG1 versus GHG2	0.093	−0.296	0.027	0.055	−0.434

[a] Geopotential heights are 850, 500, and 200 mbar, respectively. Values in parentheses exceed the 0.95 confidence level.

mbar geopotential height and temperature plates. For each variable the 0.95 significance level of the spatial correlation coefficients is estimated as $2\sigma_n$, where $\sigma_n = n^{-1/2}$ and n is the number of degrees of freedom of the trend field considered. The number of degrees of freedom is estimated as the number of empirical orthogonal functions which are necessary to explain 0.95 of the total variance of the time series of the anomalies of the variable considered. The GHGs data have been used in order to compute n for each variable; using the NCEP, GHG1, or GHG2 data does not provide significantly different estimates for n. The number of degrees of freedom turns out to be ∼50 for the annual means and ∼25 for the seasonal values for all the variables considered.

[8] For the NCEP and GHGs data sets the trend of the NH spatial mean and its 0.95 confidence interval are computed for each of the variables considered, following the procedure described by *Weatherhead et al.* [1998]. To compute the annual mean trend, a linear trend model of the form $Y_s = \mu + \omega X_s + S_s + N_s$ is assumed. The time index s runs from 1 to $m \times y$, where m is the number of records per year and y is the total number of years, 4 and 41, respectively, in our case, since we deal with seasonal data from 1960 to 2000. Y_s is the actual NH spatial mean; μ is a constant term; $X_s = s/4$ is the linear trend function; ω is the magnitude of the trend per year; S_s is the seasonal oscillating term, which is identical every fourth season; and N_s is the noise. We assume N_s to be an AR(1) noise; that is, the following relationship holds: $N_s = \phi N_{s-1} + \epsilon_s$, where $\phi = \mathrm{Corr}(N_s, N_{s-1})$ and ϵ_s is random uncorrelated noise. The seasonal term S_s is subtracted from the data set to simplify the linear trend model as $Y_s = \mu + \omega X_s + N_s$; minimizing the residual, the best estimates $\bar{\mu}$ and $\bar{\omega}$ are obtained. Defining the $\sigma(N)$ as the standard deviation of the noise N, the standard deviation of the trend $\sigma(\omega)$ is obtained as

$$\sigma(\omega) = \sigma(N)(1-\phi^2)^{1/2}\left(\frac{h_1}{h_1 h_3 - h_2^2}\right)^{1/2}, \tag{1}$$

where the h values are defined as

$$h_1 = (my-1)(1-\phi)^2 + (1-\phi^2), \tag{2}$$

$$h_2 = \frac{1-\phi}{m}\left[\frac{1}{2}my(my-1)(1-\phi) + my + \phi\right], \tag{3}$$

$$h_3 = \frac{1}{m^2}\left[\frac{1}{6}my(my+1)(2my+1)(1-\phi)^2 + m^2y^2\phi(1-\phi) + my\phi - \phi^2\right]. \tag{4}$$

If my is large, the formula for $\sigma(\omega)$ can be approximated as

$$\sigma(\omega) \approx \left(\frac{12}{m}\right)^{1/2}\frac{\sigma(N)}{y^{3/2}}\left(\frac{1+\phi}{1-\phi}\right)^{1/2}. \tag{5}$$

It should be observed that if $m = 12$, the approximate formula presented here in equation (5) coherently matches with that presented in equation (2) of *Weatherhead et al.* [1998], where the results were obtained for monthly data. The 0.95 confidence interval is defined as $[\bar{\omega} - 2\sigma(\omega), \bar{\omega} + 2\sigma(\omega)]$.

[9] To compute the seasonal NH mean trend confidence intervals, we follow the same procedure: in this case, $m = 1$; that is, there is only one record per year (the season considered), and obviously, the term S is not present. In order to

Table 4. NH Spatial Means of Seasonal and Annual Trends for 1960–2000[a]

Temperature	Data Set	Trend Statistics	DJF	MAM	JJA	SON	ANN
TAS	GHGs	$\bar{\omega}$	0.17	0.16	0.11	0.17	0.16
		$2\sigma(\omega)$	0.03	0.05	0.03	0.03	0.02
	NCEP	$\bar{\omega}$	0.12	0.14	0.11	0.11	0.12
		$2\sigma(\omega)$	0.07	0.05	0.04	0.06	0.04
T850	GHGs	$\bar{\omega}$	0.14	0.15	0.13	0.17	0.15
		$2\sigma(\omega)$	0.02	0.04	0.03	0.02	0.02
	NCEP	$\bar{\omega}$	0.14	0.20	0.18	0.16	0.17
		$2\sigma(\omega)$	0.08	0.06	0.05	0.07	0.03
	NCEP[a]	$\bar{\omega}$	0.15	0.22	0.19	0.17	0.18
		$2\sigma(\omega)$	0.08	0.06	0.05	0.08	0.03
T500	GHGs	$\bar{\omega}$	0.12	0.15	0.18	0.18	0.16
		$2\sigma(\omega)$	0.02	0.04	0.05	0.02	0.02
	NCEP	$\bar{\omega}$	0.07	0.04	0.05	0.08	0.06
		$2\sigma(\omega)$	0.08	0.06	0.07	0.07	0.04
T200	GHGs	$\bar{\omega}$	0.12	0.13	0.19	0.19	0.16
		$2\sigma(\omega)$	0.03	0.05	0.04	0.03	0.02
	NCEP	$\bar{\omega}$	0.14	0.05	0.05	0.11	0.09
		$2\sigma(\omega)$	0.11	0.15	0.16	0.14	0.08

[a] Temperature abbreviations are surface air, 850, 500, and 200 mbar, respectively. Half widths are for the 0.95 confidence intervals. These results have been obtained excluding the grid cells where T850 is not defined in the model. Trends are in K/decade.

Table 5. NH Spatial Means of Seasonal and Annual Trends for 1960–2000 for Surface Pressure[a]

Data Sets	Trend Statistics	DJF	MAM	JJA	SON	ANN
GHGs	$\bar{\omega}$	0.00	−0.01	−0.01	−0.03	−0.01
	$2\sigma(\omega)$	0.05	0.03	0.04	0.04	0.03
NCEP	$\bar{\omega}$	0.06	0.12	0.23	0.18	0.15
	$2\sigma(\omega)$	0.06	0.10	0.10	0.10	0.04

[a] Trends are in mbar/decade. Half widths are for the 0.95 confidence intervals.

obtain more homogeneous results the NCEP NH mean trends for 850 mbar temperature and geopotential height have been computed both considering and excluding the grid cells where those variables are not defined in the GHGs data sets. The difference between these two techniques is fairly small.

[10] It should be stressed that the procedure of using all the seasonal values instead of the annual average in order to compute the annual trend reduces the 0.95 confidence interval of such a trend (while essentially retaining the same best estimate), the fundamental reason being that more information is retrieved, since we use $4y$ data instead of y. An important consequence is to have smaller confidence intervals for the annual trends than for the seasonal ones. This fact can be particularly important for those data sets, like MSU, which have such a short time length that trends can hardly be recognized. Applying this procedure to already analyzed data, like the *Angell* [1999] radiosonde records, it is possible to obtain better constrained statistics than those published. A detailed comparison of NCEP reanalysis data, Angell radiosonde data, and MSU data will be presented elsewhere (V. Lucarini, unpublished results, 2001).

3. Results

[11] Tables 1, 2, and 3 present the NH spatial correlation analysis for temperature, surface pressure, and geopotential height, respectively. The results that are distinguishable from noise with a 0.95 confidence level are underlined. Tables 4–6 present NH mean decadal trends and their 0.95 confidence intervals for the same quantities. Figures 1–8 give a portrait of the spatial patterns of GHGs and NCEP winter (DJF) and annual (Ann) 41 year trends of the variables analyzed. For most variables considered, the highest correlations between NCEP data and model experiments occur in winter and spring (with the notable exception of the surface temperature that has maximum correlation in fall), while the poorest agreement occurs in summer. One explanation is that the model has deficiencies in some aspects of climate that play a more relevant role in summer, like cloud feedback and the hydrological cycle. In addition, summer data are more noisy, as can be seen from the fact that the confidence intervals of the mean trends are much wider in summer for all variables for both GHGs and NCEP data sets. The spatial correlation of the annual means is usually fairly good and resembles the corresponding winter correlation thanks to the fact that the winter signal usually has the strongest local features and at the same time has the least noise. It should be observed that for every variable the spatial correlation between the two model experiments is smaller than those between each of them and the NCEP data. One reason for this could be the presence of different climate drifts in the two subtracted control simulations mentioned in section 2. This relatively poor agreement between GHG1 and GHG2 might also suggest that the natural variability of the model is larger than the observed variability. Another general characteristic of the presented results is that the GHGs trends usually underestimate the local maxima and minima compared with NCEP. This is partly beacause GHGs is an average of two experiments, while NCEP is a single realization. Another possible explanation is that this model presents an Arctic Oscillation (AO) index trend which is about one third of the observed value [*Shindell et al.*, 1999], thus having a smaller increase of the average intensity of western winds over the Atlantic Ocean. Since the presence of an AO index trend essentially creates a zonal redistribution of heat, its underestimation should not dramatically effect the reliability of mean hemispheric trends.

[12] The NH mean trend confidence intervals match well for all variables for both seasonal trends and the annual trend, with the interesting exception of surface pressure and 850 mbar geopotential height. The widths of the confidence intervals determined from the GHGs data are usually smaller than those deduced from NCEP data; the latter become larger than the former as we look higher in the atmosphere, where NCEP data seem very noisy. It is important to stress that the small widths of the confidence intervals of the trends deduced from the GHGs data sets are related to the procedure through which these data sets were created; a GHGs confidence interval should generally be smaller than either the GHG1 or GHG2 confidence interval itself. We observe that in agreement with the theory presented in section 2, the confidence interval for annual trends is always smaller than those for seasonal trends.

[13] While surface air temperature is dominated by the radiation budget, 850 mbar fields are strongly affected by the hydrologic cycle and topography. Model deficiencies in handling these effects contribute to poorer comparisons with NCEP for the 850 mbar fields. We now present, variable by variable, some comments about the spatial correlations and

Table 6. NH Spatial Means of Seasonal and Annual Trends for 1960–2000 for Geopotential Height[a]

Geopotential Height	Data Set	Trend Statistics	DJF	MAM	JJA	SON	ANN
Z850	GHGs	$\bar{\omega}$	0.7	0.6	0.5	0.4	0.6
		$2\sigma(\omega)$	0.4	0.3	0.3	0.4	0.2
	NCEP	$\bar{\omega}$	1.3	2.0	3.0	2.4	2.2
		$2\sigma(\omega)$	0.7	0.6	1.1	0.9	0.5
	NCEP[d]	$\bar{\omega}$	1.2	1.8	2.7	2.2	2.0
		$2\sigma(\omega)$	0.7	0.6	1.1	1.0	0.5
Z500	GHGs	$\bar{\omega}$	2.6	2.9	2.9	3.2	2.9
		$2\sigma(\omega)$	0.4	0.9	0.5	0.5	0.3
	NCEP	$\bar{\omega}$	3.1	3.6	4.5	4.1	3.9
		$2\sigma(\omega)$	1.8	1.2	1.7	1.2	0.9
Z200	GHGs	$\bar{\omega}$	6.2	6.9	8.4	8.6	7.7
		$2\sigma(\omega)$	0.8	2.0	1.6	1.0	0.8
	NCEP	$\bar{\omega}$	6.1	5.1	6.6	7.2	6.2
		$2\sigma(\omega)$	4.6	3.6	4.4	3.8	2.2

[a] Trends are in m/decade. Geopotential heights are 850, 500, and 200 mbar, respectively. Half widths are for the 0.95 confidence intervals and have been obtained excluding the grid cells where Z850 is not defined in the model.

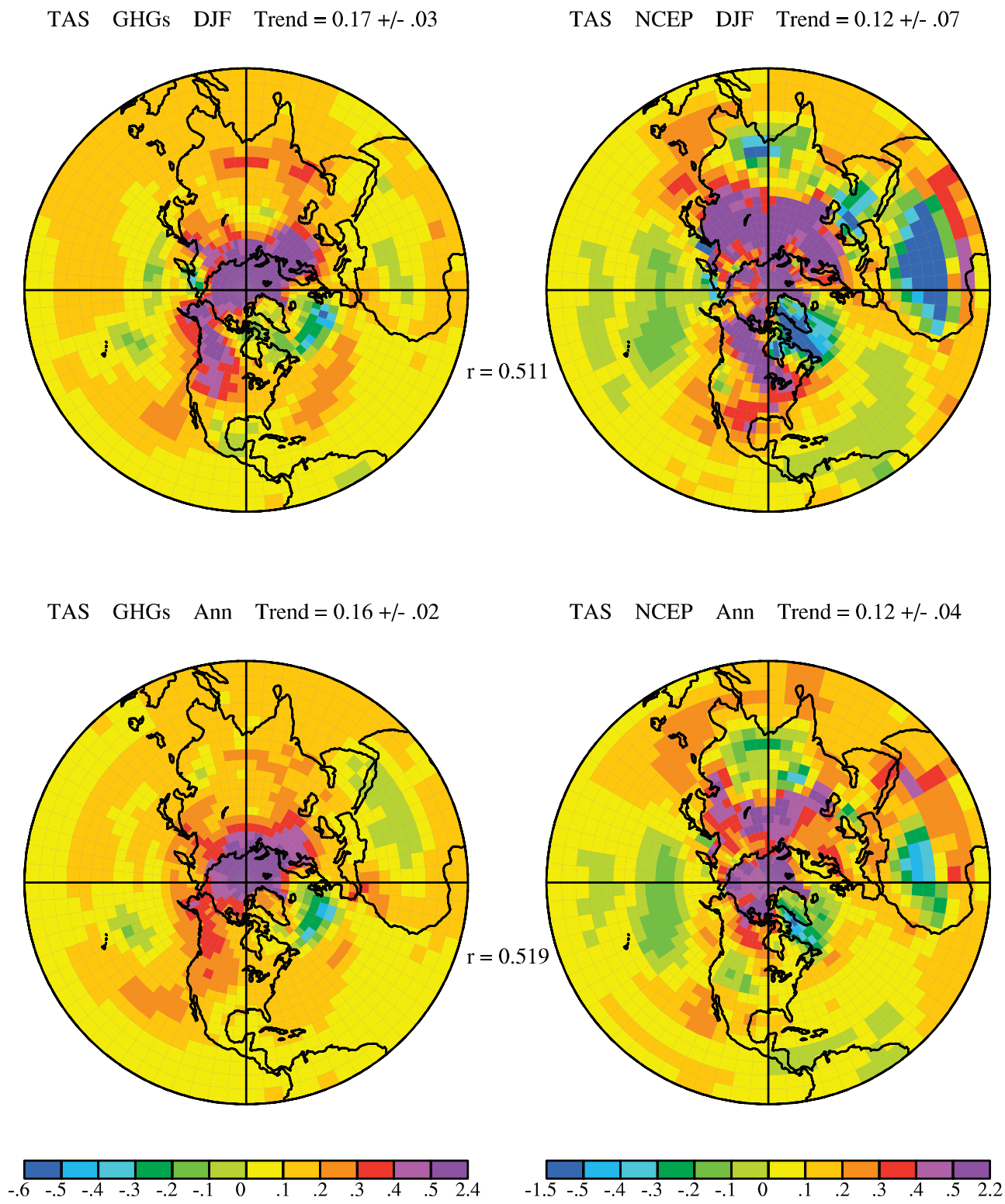

Figure 1. Spatial patterns of surface air temperature trends (K/decade) for data sets GHGs and NCEP, for winter (DJF) and annual (Ann), and for years 1960–2000. "Trend" is the Northern Hemisphere mean trend plus or minus the half width of the 0.95 confidence interval, and r is the spatial correlation coefficient between adjacent fields.

the NH mean trends obtained from NCEP and GHGs data sets.

[14] The main disagreement between the two trend maps showing surface temperature (Figure 1) is that in the GHGs data, Greenland, the western United States, and southern Asia are not cooling and the cooling in the Sahara desert and the heating in central Siberia are underestimated. The spatial patterns between the model and NCEP match well

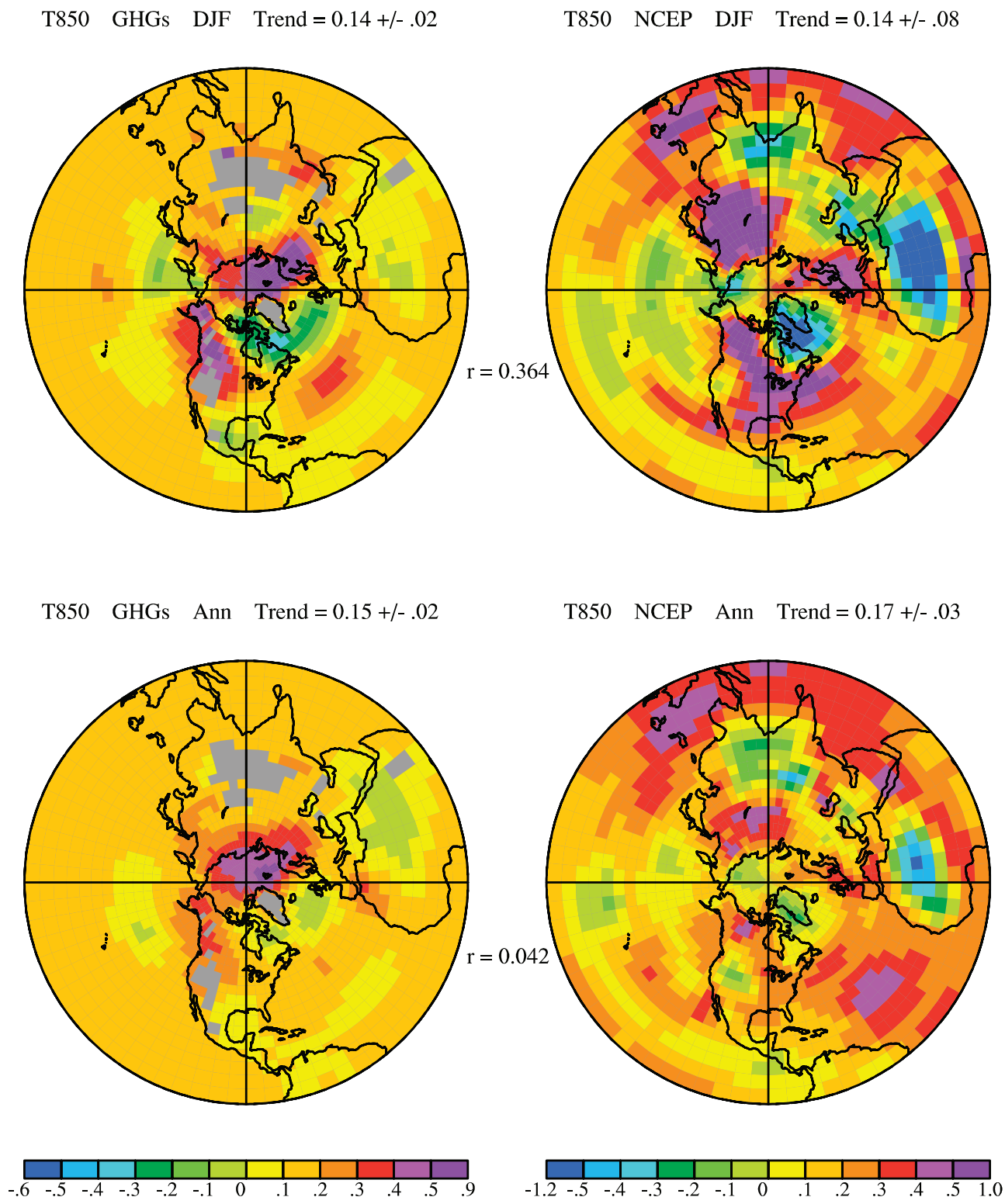

Figure 2. Same as Figure 1, except for 850 mbar temperature trends.

over the oceans, with warming over the Arctic Ocean and cooling over the North Atlantic [*Russell and Rind*, 1999] and North Pacific Oceans. The overall effect of the disagreement is the presence of slightly higher mean trends for the GHGs data, although the confidence intervals still overlap.

[15] The model results for temperature at 850 mbar (Figure 2) agree with NCEP in the western portion of the hemisphere, while the patterns do not match in the eastern portion, particularly for the annual trend. Except for winter and spring the spatial correlation coefficients between the model and NCEP are small or negative. The 850 mbar

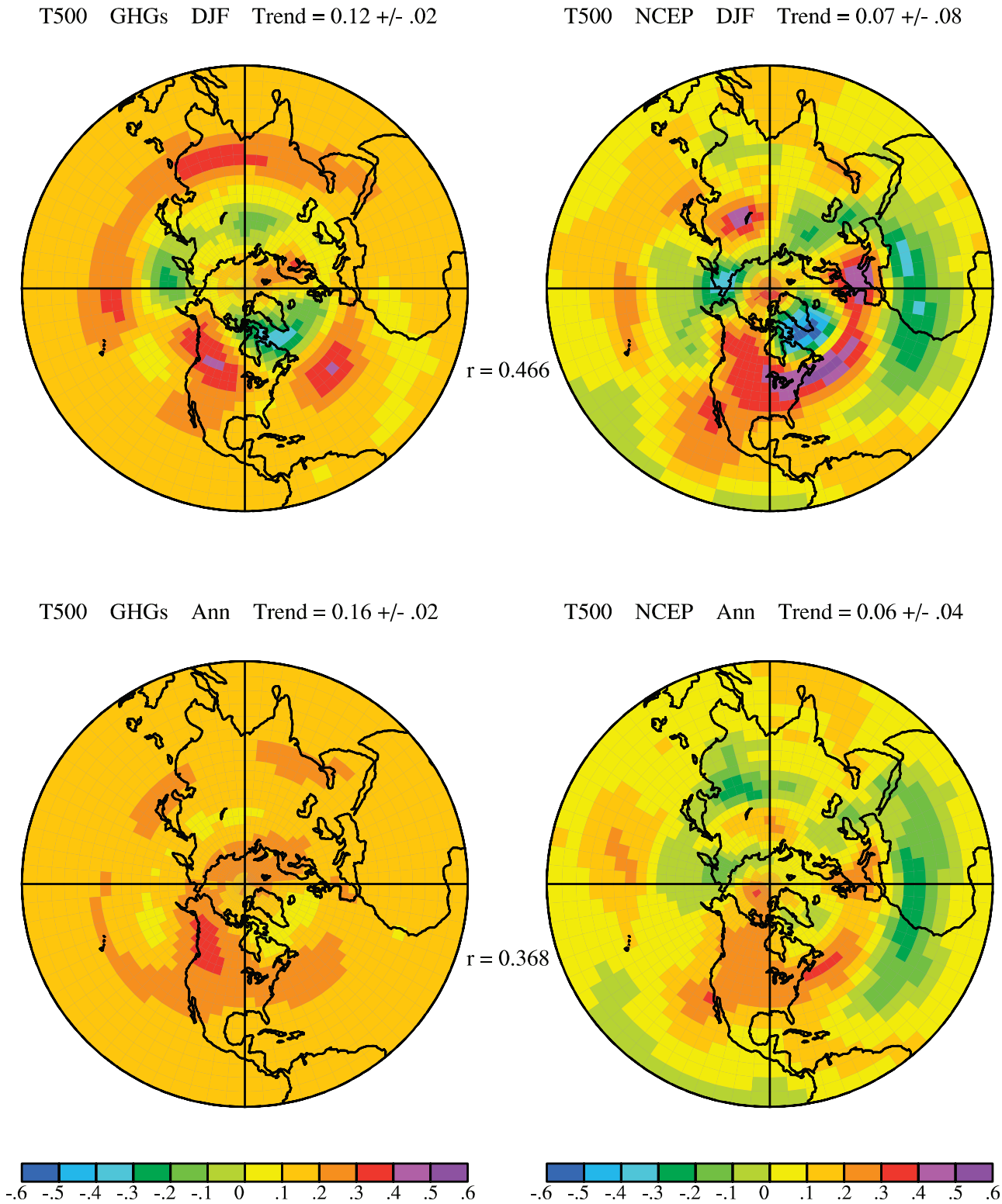

Figure 3. Same as Figure 1, except for 500 mbar temperature trends.

temperature over the Arctic Ocean is warming in the model (as it is at the surface), whereas NCEP shows a small trend on the annual average. On the contrary, the hemispheric mean trends match very well for all seasons and for the annual trends, presenting strongly overlapping confidence intervals.

[16] For temperature at 500 mbar (Figure 3) the agreement over the northwestern quarter sphere is excellent. The model, however, misses some NCEP extremes and is too hot in the northeast, so that a large bias between the NCEP and GHGs hemispheric mean trends exists.

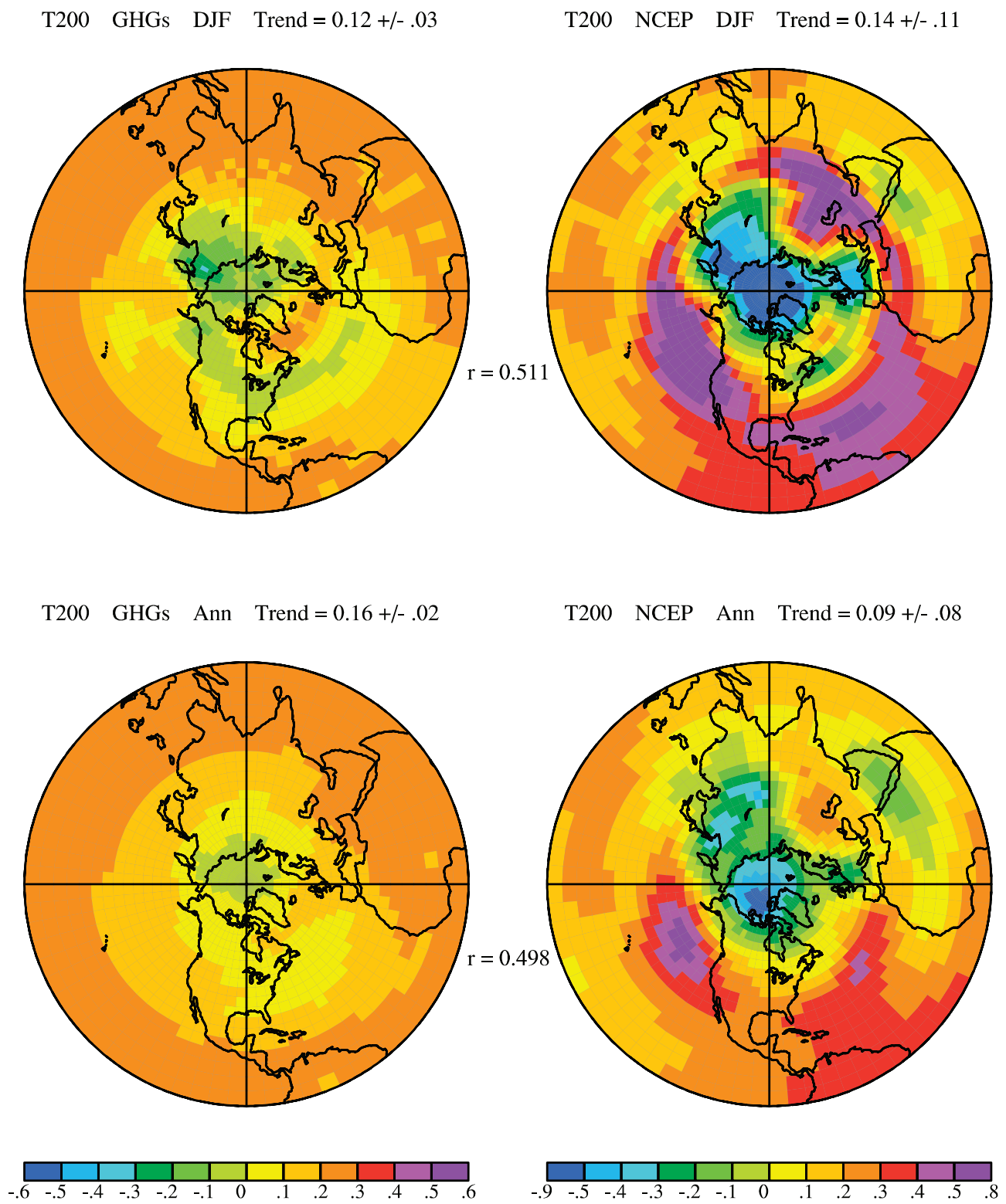

Figure 4. Same as Figure 1, except for 200 mbar temperature trends.

[17] The GHGs and NCEP trend patterns for temperature at 200 mbar (Figure 4) are similar, presenting a northward decrease of the trends over the Arctic Ocean which is opposite to that which occurs at the surface. The NCEP data show a very deep minimum over the Arctic Ocean, the intensity of which is not captured by GHGs data. This causes a large difference between the NH mean trends except in winter; the model is always too hot.

[18] The model for surface pressure (Figure 5) captures the spatial patterns of the NCEP observations. In particular, there is a signal of an increase in the AO index over time [*Shindell et al.*, 1999]. The hemispheric mean trend

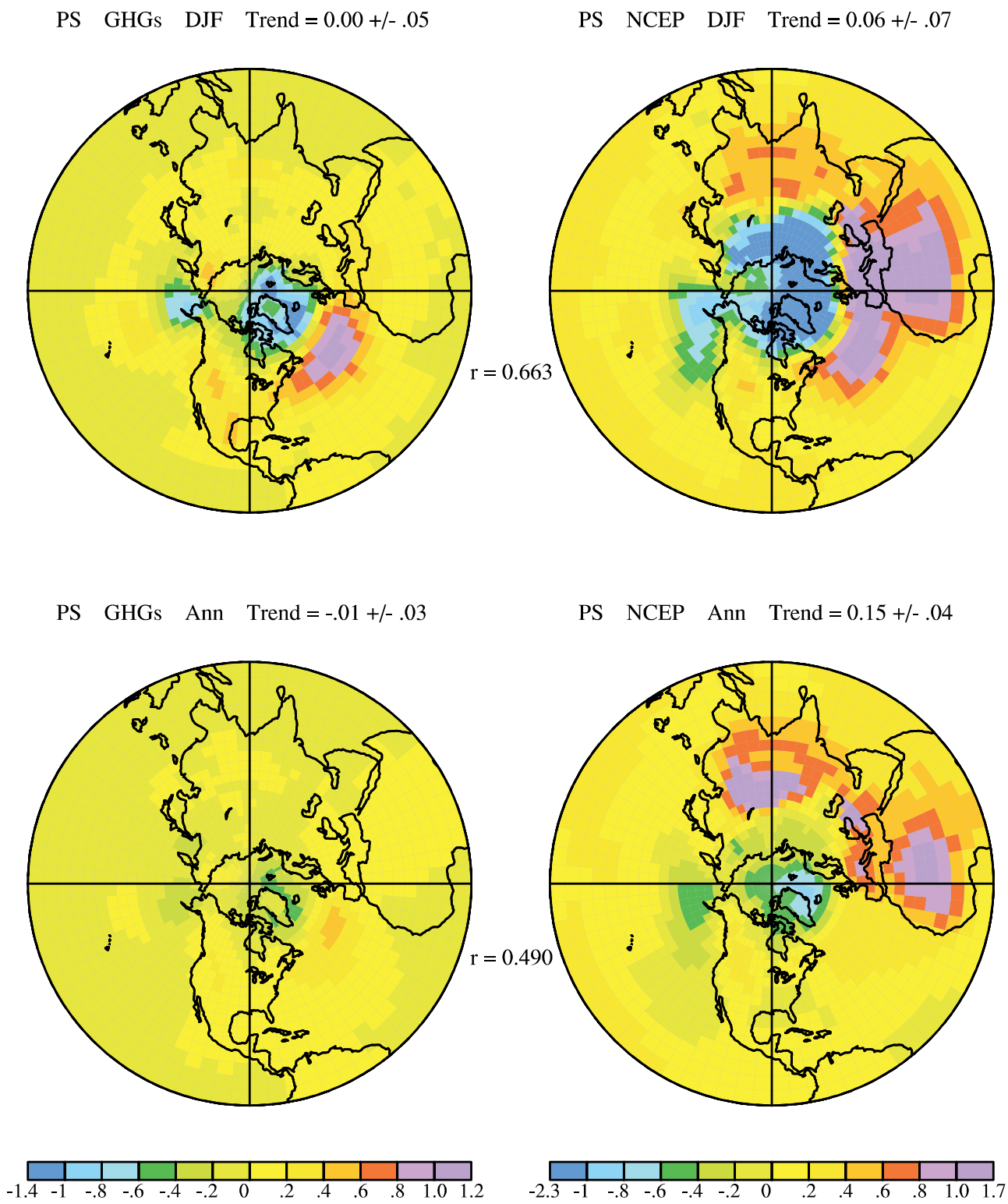

Figure 5. Spatial patterns of surface pressure trends (mbar/decade) for GHGs and NCEP, for winter (DJF) and annual (Ann), and for years 1960–2000. "Trend" is the NH mean trend plus or minus the half width of the 0.95 confidence interval, and *r* is the spatial correlation coefficient between adjacent fields.

of the model is negligible and not compatible with NCEP's trend, which is strongly positive, the main cause being an underestimation of the positive trends over Europe, Africa, and southern Asia. The model's surface pressure presented here is the dry atmospheric pressure which is globally constant. The model does indicate an increase of global surface pressure of ~0.02 mbar/decade due to the increase of humidity between 1960 and 2000. The NCEP data indicate a more significant shift of mass from the SH into the NH.

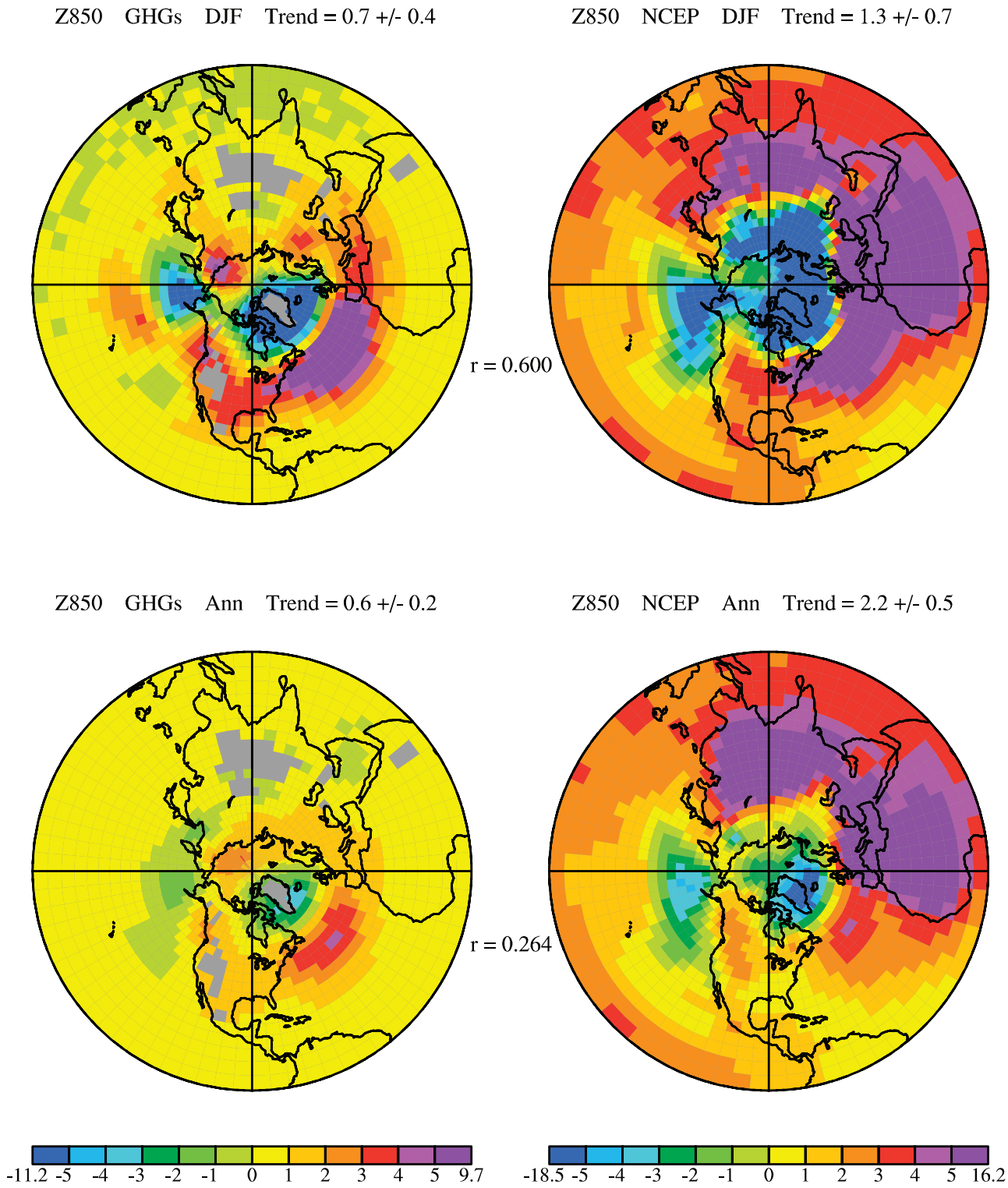

Figure 6. Spatial patterns for 850 mbar geopotential height trends (m/decade) for GHGs and NCEP, for winter (DJF) and annual (Ann), and for years 1960–2000. "Trend" is the NH mean trend plus or minus the half width of the 0.95 confidence interval, and r is the spatial correlation coefficient between adjacent fields.

[19] The model for geopotential height at 850 mbar (Figure 6) and NCEP compare favorably over the oceans and North America but are poorly correlated over Asia. The model's mean trends are considerably weaker than those of NCEP, and the confidence intervals, in general, do not overlap.

[20] For geopotential height at 500 mbar (Figure 7) the patterns in the western part of the NH present good agree-

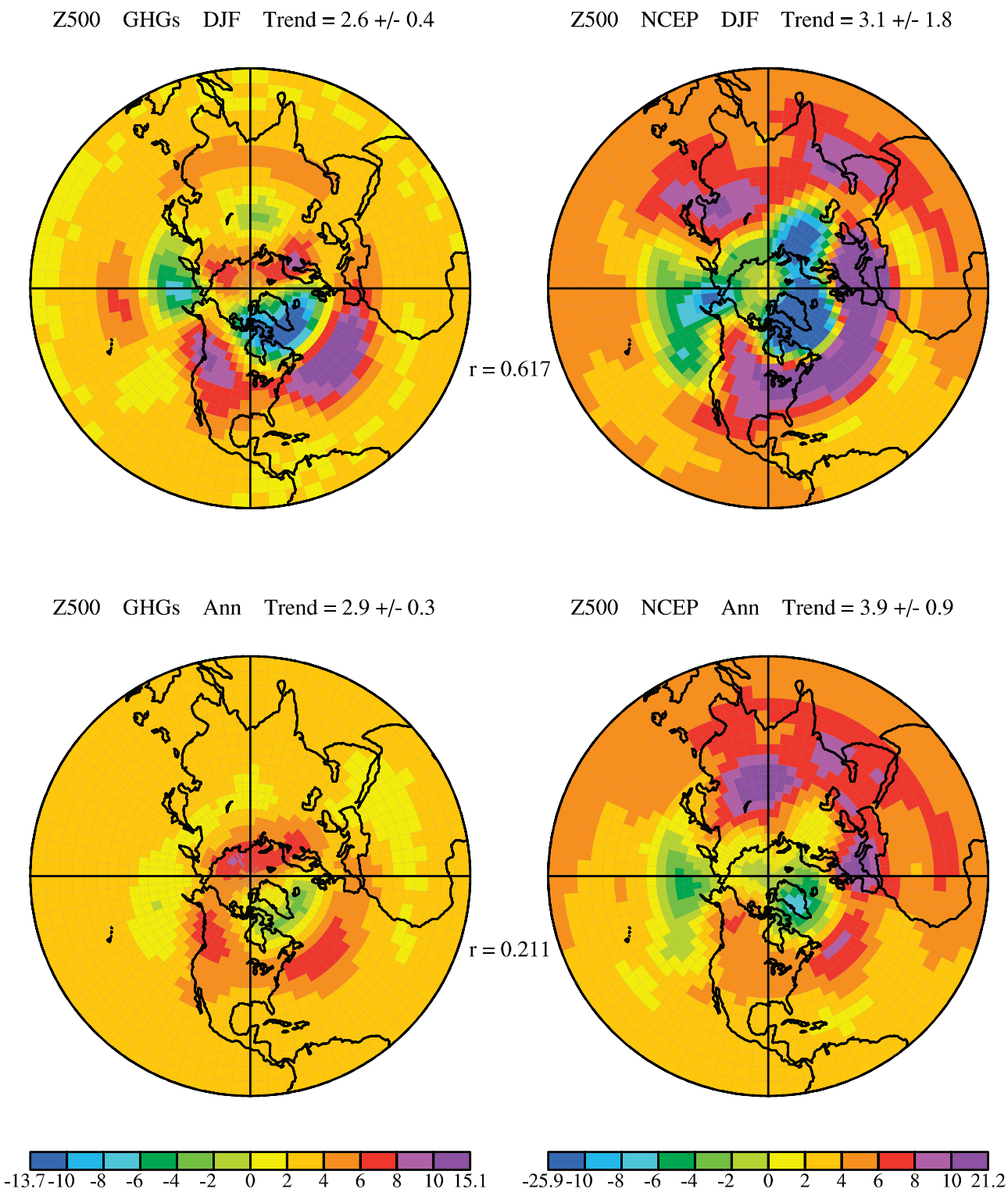

Figure 7. Same as Figure 6, except for 500 mbar geopotential height trends (m/decade).

ment in winter and in the annual average, while the GHGs trends in the eastern part are poorly correlated with NCEP's. The model underestimates the mean trends because it underestimates the maxima more seriously. In any case, the confidence intervals barely overlap.

[21] The NCEP data for geopotential height at 200 mbar (Figure 8) feature deeper minima and higher maxima. As seen for the 200 mbar temperature, the NCEP confidence intervals are much greater than those of the model, which indicates a much greater interannual variability of 200 mbar quantities.

4. Conclusions

[22] A thorough comparison between NCEP reanalysis data and two GISS AOM greenhouse gases forced runs

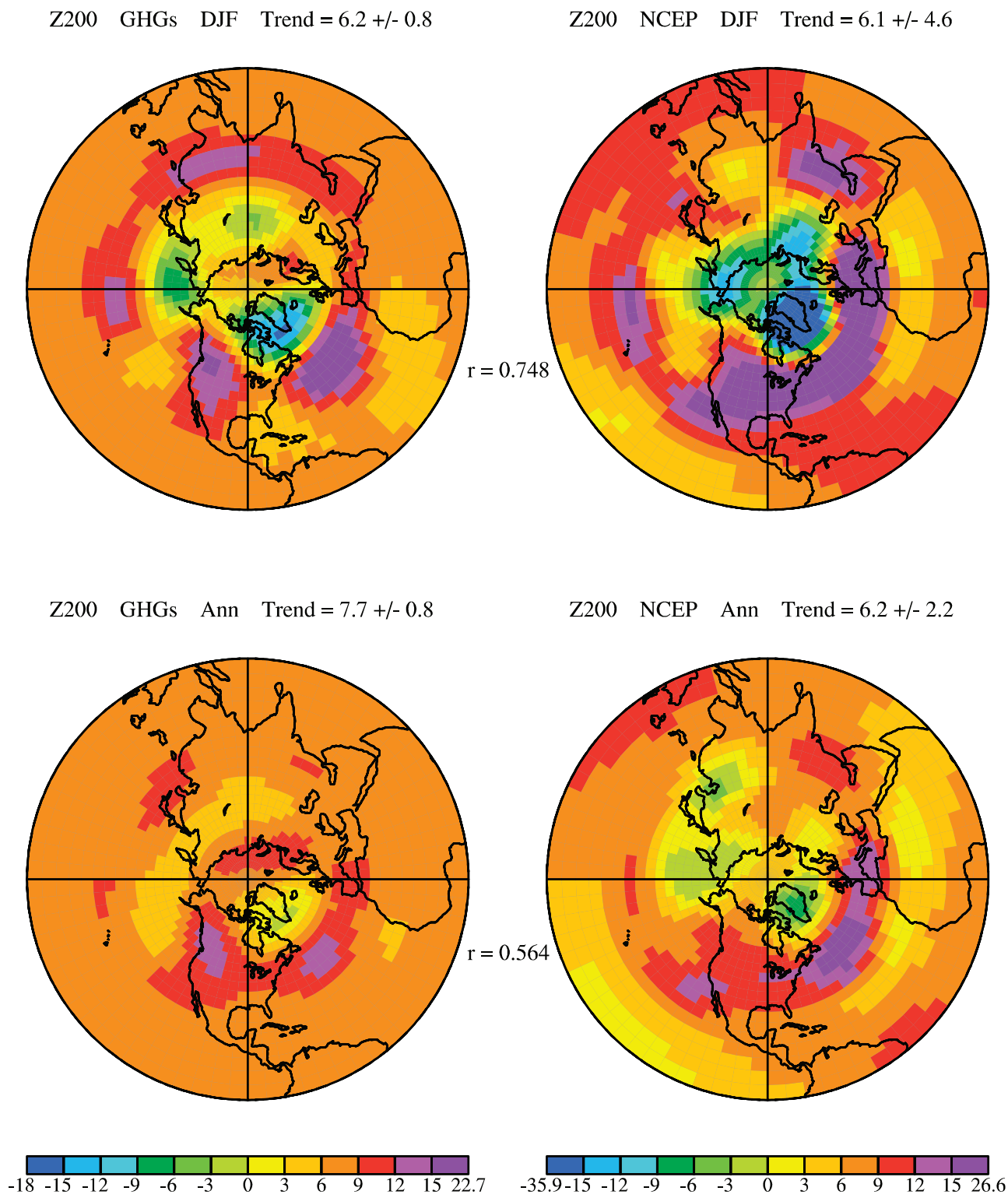

Figure 8. Same as Figure 6, except for 200 mbar geopotential height trends (m/decade).

have been performed in order to assess the ability of this model to describe regional and NH mean trends for several climatologically relevant variables: surface pressure and temperature and geopotential height and temperature at 850, 500, and 200 mbar. A spatial correlation analysis has been performed with statistically significant positive results for winter and annual trends for most variables considered, while the agreement is very poor for the 850 mbar temperature. The NH mean trends, together with their 0.95 confidence intervals, have been computed, showing for all variables, except surface pressure and geopotential height at 850 mbar, a good agreement between the NCEP and GISS AOM outputs in terms of statistical significance for most seasonal and annual trends. Computing the annual

trends using all the seasonal data instead of the annual average is statistically more efficient because more information is retrieved, thus reducing the confidence intervals. The GISS AOM has been able to capture the climatological evolution of the last 40 years as described by NCEP with good accuracy in describing both local features and NH average trends; this suggests that this model may be reliable for future projections. This study more generally stresses the importance of using mathematical tools to capture the compatibility of both regional and global results of trends deduced from a model and observations in order to assess a model's reliability and efficacy in forecasting future climate change more rigorously. More information and quantities for these AOM simulations are available at http://aom.giss.nasa.gov.

[23] **Acknowledgments.** We are grateful to P. Stone, D. Rind, and G. Schmidt for useful suggestions. One author (V. L.) would like to thank G. Schmidt, D. Rind, and R. Miller for stimulating conversations and all of GISS for the warm hospitality received while visiting during the summer of 2001.

References

Angell, J. K., Comparison of surface and tropospheric temperature trends estimated from a 63-station radiosonde network, 1958–1998, *Geophys. Res. Lett.*, *26*, 2761–2764, 1999.

Angell, J. K., Global, hemispheric, and zonal temperature deviations derived from radiosonde records, in *Trends Online: A Compendium of Data on Global Change*, Carbon Dioxide Inf. Anal. Cent., Oak Ridge, Tenn., 2000. (Available at http://cdiac.esd.ornl.gov/trends/temp/angell/angell.html).

Basist, A. N., and M. Chelliah, Comparison of temperatures derived from the NCEP/NCAR reanalysis, NCEP operational analysis, and the microwave sounding unit, *Bull. Am. Meteorol. Soc.*, *78*, 1431–1447, 1997.

Chelliah, M., and C. F. Ropelewski, Reanalysis-based tropospheric temperature estimates: Uncertainties in the context of global climate change detection, *J. Clim.*, *13*, 3187–3207, 2000.

Gaffen, D. J., M. A. Sargent, R. E. Habermann, and J. R. Lanzante, Sensitivity of tropospheric and stratospheric temperature trends to radiosonde data quality, *J. Clim.*, *13*, 1776–1796, 2000.

Hansen, J. E., R. Ruedy, and M. Sato, Global surface air temperature in 1995: Return to pre-Pinatubo level, *Geophys. Res. Lett.*, *23*, 1665–1668, 1996.

Hansen, J. E., R. Ruedy, J. Glascoe, and M. Sato, GISS analysis of surface temperature change, *J. Geophys. Res.*, *104*, 30,997–31,022, 1999.

Hurrell, J. W., and K. E. Trenberth, Difficulties in obtaining reliable temperature trends: Reconciling the surface and satellite microwave sounding unit records, *J. Clim.*, *11*, 945–967, 1998.

Hurrell, J. W., S. J. Brown, K. E. Trenberth, and J. R. Christy, Comparison of tropospheric temperatures from radiosondes and satellites: 1979–98, *Bull. Am. Meteorol. Soc.*, *81*, 2166–2177, 2000.

Johns, T. C., J. M. Gregory, P. A. Stott, and J. F. B. Mitchell, Correlations between patterns of 19th and 20th century surface temperature change and HadCM2 climate model ensembles, *Geophys. Res. Lett.*, *28*, 1007–1010, 2001.

Kalnay, E., et al., The NCEP/NCAR 40-Year Reanalysis Project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471, 1996.

Levitus, S., R. Burgett, and T. P. Boyer, *World Ocean Atlas 1994*, Natl. Oceanic and Atmos. Admin., Silver Spring, Md., 1994.

Russell, G. L., and D. Rind, Response to CO^2 transient increase in the GISS coupled model: Regional cooling in a warming climate, *J. Clim.*, *12*, 531–539, 1999.

Russell, G. L., J. R. Miller, and D. Rind, A coupled atmosphere-ocean model for transient climate change studies, *Atmos. Ocean*, *33*, 683–730, 1995.

Russell, G. L., J. R. Miller, D. Rind, R. A. Ruedy, G. A. Schmidt, and S. Sheth, Comparison of model and observed regional temperature changes during the last 40 years, *J. Geophys. Res.*, *105*, 14,891–14,898, 2000.

Shindell, D. T., R. L. Miller, G. A. Schmidt, and L. Pandolfo, Simulations of recent northern winter climate trends by greenhouse gas forcing, *Nature*, *339*, 569–572, 1999.

Trenberth, K. E., D. P. Stepaniak, J. W. Hurrell, and M. Fiorino, Quality of reanalyses in the tropics, *J. Clim.*, *14*, 1499–1510, 2001.

Weatherhead, E. C., et al., Factors affecting the detection of trends: Statistical considerations and applications to environmental data, *J. Geophys. Res.*, *103*, 17,149–17,161, 1998.

V. Lucarini, Department of Earth, Atmospheric, and Planetary Sciences, Building 54, Room 17–19, Massachusetts Institute of Technology, Cambridge, MA 02138, USA. (lucarini@mit.edu)

G. L. Russell, NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, USA. (grussell@giss.nasa.gov)